DOCUMENT RESUME

ED 266 173 TM 860 117

AUTHOR Aschbacher, Pamela; Burry, James
TITLE A Process for Designing and Implementing a Dual Purpose Evaluation System. CSE Resource Paper No. 7.
INSTITUTION California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Nov 85
GRANT NIE-G-83-0001
NOTE 40p.; In: "Resource Papers and Technical Reports. Research into Practice Project" (TM 860 116).
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Accountability; Data Collection; *Educational Assessment; Elementary Secondary Education; *Evaluation Methods; *Evaluation Needs; Information Needs; *Models; Policy Formation; *Program Evaluation; Questionnaires
IDENTIFIERS *Center for the Study of Evaluation CA; Site Visits; Task Force Approach

ABSTRACT

This paper offers a guide for a dual purpose evaluation plan that can provide information for local program managers and simultaneously serve the accountability and reporting needs of a state legislature, district office, or other policy making body. The dual focus approach has two advantages: (1) it provides a common information base for decisions at both ends of the continuum; and (2) it helps ensure that these decisions are in harmony with each other. In addition, this approach conserves time and effort of staff and students alike by making one data collection effort serve two purposes. This user oriented model reflects many years of evaluation experience by the Center for the Study of Evaluation (CSE). The model evaluation plan consists of eight steps: (1) identify evaluation users and their needs; (2) form a task force; (3) reach consensus on information needs; (4) specify information base and develop measures; (5) develop procedures for data collection and report of results; (6) pilot and refine measures and procedures; (7) collect information; and (8) prepare reports. While this plan was based on evaluation experiences in education, it is intended to be applicable to other fields also. Appendices include a form for description of programs and participants, a form for open-ended questionnaire, a sample checklist for the site visit team leader, and a sample site visit schedule. (LMO)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED266173

A PROCESS FOR DESIGNING AND IMPLEMENTING A DUAL PURPOSE EVALUATION SYSTEM

CSE RESOURCE PAPER NO. 7

1985

Pamela Aschbacher, Center for the Study of Evaluation
University of California, Los Angeles

James Burry, Center for the Study of Evaluation
University of California, Los Angeles

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

✓ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

C. Griffith

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CENTER FOR THE STUDY OF EVALUATION
Graduate School of Education
University of California, Los Angeles

TM 860 117

TABLE OF CONTENTS

	PAGE
INTRODUCTION	1
Purpose and Approach	1
Potential Users	3
Structure of the Resource Paper	4
DUAL PURPOSE EVALUATION: A MODEL PROCESS	
Step 1. Identify Evaluation Users and Their Needs	5
Step 2. Form Task Force	6
Step 3. Reach Consensus on Information Needs	10
Step 4. Specify Information Base and Develop Measures	12
Step 5. Develop Procedures for Data Collection & Report of Results	16
Step 6. Pilot and Refine Measures and Procedures	21
Step 7. Collect Information	23
Step 8. Prepare Reports	23
SUMMARY OF CRITICAL CONCERNS	24
CONCLUSION	26
BIBLIOGRAPHY	28
APPENDIX A: Example of Standard Form for Description of Program and Participants	
APPENDIX B: Example of Open-Ended Questionnaire for Qualitative Data	
APPENDIX C: Sample Checklist for Site Visit Team Leader	
APPENDIX D: Sample Site Visit Schedule	

INTRODUCTION

Purpose and Approach

This resource paper offers a guide for a dual purpose evaluation plan that can provide formative information for local program managers and simultaneously serve the accountability and reporting needs of a state legislature, district office, or other policymaking body. While the primary audience for this paper is expected to be those interested in educational evaluation, and the examples provided will be from this field, we hope that this paper will also prove interesting to those involved in evaluation of non-educational programs. The approach presented here reflects findings from several years of CSE research on a formative orientation to evaluation, on qualitative methodology, and on strategies for maximizing the utilization of evaluation efforts.

There are three principles that guide our approach. First and probably most important, we believe that evaluation should serve the needs of a multiplicity of users. Teachers, counselors, and program directors continually want evaluation information to refine and improve the program with which they work. In addition, policymakers for such programs (be they at the district, state or federal level) also need information to guide their policy- and decisionmaking. It is clear that these two demands---the top-down demand for broad-level accountability (to improve management and to elevate standards of excellence) and the bottom-up demand for adaptive, sensitive infor ation to be useful at the local level--- push in different and not totally compatible directions. Because it has usually been assumed that these different decisions require different types of data, separate

evaluations have usually been conducted to meet these two needs. If we consider the result when these evaluations are added to all the other data collection efforts that occur (e.g., for auditing, accreditation, diagnosis and prescription), we can begin to comprehend why many students, teachers and program administrators feel overwhelmed by evaluation. Furthermore, much of the data collected is of limited utility for them because it often fails to reflect the actual school context and curriculum. They view evaluation as burdensome and intrusive and come to resent the top-down demand for accountability, seeing it as more of a liability than an asset.

Other problems arise when policymakers and program staff have no common basis for their separate inferences about policies and practices. irst, the general intent of educational policy formation is to improve the quality of educational services and to help students attain the highest levels of compentency in school subjects. To accomplish this goal, the policies must be translated into practices that are compatible with the views, needs and capabilities of teachers and students. Second, local programs alone do not necessarily have sufficient resources to solve their problems. The solutions may require initiative, directions, resources and action at those higher levels charged with responsibility for governance, resource allocation and policy formation. In both cases there is high potential for slippage when the information used to assess quality and formulate policy functions independently from that used to actually teach students.

In summary, the current system of independent evaluations appears both uneconomical of time and effort and ineffective in accomplishing the goals

of either end of the educational continuum. A partial solution to this problem, presented here, is a single evaluation effort that uses a common information base to serve both policy and program needs simultaneously.

The second principle guiding our development of this paper is that evaluation, testing and standard setting are endeavors which are partly technical, partly political, and partly social. Technical expertise is essential in measurement development and analysis, to ensure valid and reliable use of results. Social understanding is essential to ensure fairness and utility. Similarly, evaluation questions arise out of people's information requirements, while the design and interpretation of evaluations depend on technical competence. The definition of standards depends on values and consensus; the measurement of their attainment involves technical considerations. Thus it seems crucial that a useful evaluation plan will draw on all these areas of expertise and involve professional evaluators as well as the range of users at both levels of the system.

The final principle is that effective evaluation plans are characterized by several key features:

(a) evaluators who are personally committed to the use of results and who have political sensitivity, credibility, and rapport with users;

(b) users who are also committed to implementing the results of the valuation, open to change, and involved in the process so that it reflects their leadership, expectations, and perceptions of needs and risks;

(c) a setting in which representatives of the major stakeholders agree on the focus of the evaluation and the kinds of information it should

produce, and in which local program personnel have the autonomy to act upon the evaluation's findings; and

(d) the evaluation procedure is decision-based, allows purposeful sharing of ideas, and provides timely, specific and relevant information (Burry, Alkin & Ruskus, 1985).

Potential Users

The evaluation process described in this paper reflects the need for a systematized, short term, qualitative evaluation plan that addresses both the desire of local programs for formative, program planning and improvement-oriented information as well as the needs of policymakers for standardized information across multiple sites for the purpose of summative decisionmaking, often officially mandated. This evaluation process is intended to be appropriate in a setting characterized by multiple program sites with similar missions (although a single program site could also make use of this plan) overseen by a separate policymaking body. Such settings might include a foundation or community agency with several programs to oversee or the typical educational settings at every age level and level of organization. Specific examples include: (a) a state office overseeing special admissions programs for minority students at each campus of a statewide community college system, (b) a county agency overseeing child care information and referal systems in several cities, and (c) a school district office overseeing bilingual programs in the elementary schools in the district.

Structure of the Resource Paper

The dual purpose evaluation model presented here is described in eight steps:

1. Identify evaluation users and their needs.
2. Form task force.
3. Reach consensus on information needs.
4. Specify common information base and develop measures.
5. Develop procedures.
6. Pilot and refine measures and procedures.
7. Collect data.
8. Prepare reports for policymakers and program personnel.

Where appropriate, specific examples are provided from our experiences with several evaluation projects. The report concludes with a summary of critical concerns.

DUAL PURPOSE EVALUATION PLAN: A MODEL PROCESS

Step 1. Identify Evaluation Users & Their Needs

This evaluation process is designed to meet the needs of two levels of users: policymakers and program personnel. Each of these has somewhat different needs. The purpose of this first step is to identify the specific policymakers (e.g. members of the school board and state legislators) who are responsible for the programs and to identify those individuals involved on-site with the actual implementation of a program (e.g. faculty, staff, and local administrators).

The policymakers usually need summative indicators reflecting program goals and outcomes across all sites at which a particular type of program is implemented. These indicators will focus on broad patterns and trends emerging across program sites rather than on the program as implemented at an individual site. To obtain this information policymakers need a

comparable conceptual scheme and measurement base for each site. Where many sites are involved, they may employ an evaluation cycle in which some percent of the total number of sites is evaluated in a given year and each site participates only once every several years. Thus in a given year, the data is collected from only a sample of the whole system and is aggregated so as to provide a view of the system rather than of the individual program sites monitored that year.

The program personnel, on the other hand, need specific information to pinpoint the strengths and weaknesses of their individual implementation of the program in order to make improvements. They want to assess the extent to which the intended program has actually been implemented and the outcomes that have been achieved, both intended and unintended.

Step 2. Form Task Force

Including representatives of all major users of the evaluation (at both the program and policymaking levels) as members of a task force to oversee the evaluation facilitates both the process and the utilization of results. Many an evaluation has been undermined by evaluation users who failed to understand the intent, felt threatened by the potential for change, had political agendas at variance with the goals of the evaluation, were used to pro forma evaluations in which results were never utilized, or were faced with information that could not be clearly interpretted for action. Involving users early in the process helps avoid these problems. It provides valuable input at the point when it is most useful rather than later, thereby causing duplication of effort. It also gives those participants a sense of ownership of the evaluation which makes it much

more likely that the results will actually be used for the intended purposes.

The responsibilities of a task force are:

(a) to define the dual focus of the evaluation and reach consensus on what information is needed in the evaluation,

(b) precisely specify the information base and select measures to collect the information,

(c) plan the data collection and analysis procedures,

(d) pilot test and refine the measures, materials, and procedures,

(e) oversee collection of the information,and

(f) prepare reports for policymakers and for program personnel.

In some cases, the task force may preside over a single evaluation. In other circumstances, the task force may be an ongoing body (whose members may change over time) that oversees all two-tiered evaluations of the target programs. In either case, the task force will consist of the general task force members, the task force director, and one or more evaluation experts.

What kinds of backgrounds and skills are sought in task force members? First, they need to represent the major stakeholders in the evaluation: policymakers and program personnel. Where programs serve a number of special groups or provide several services (e.g. a program for the handicapped may serve those with learning disabilities and those with visual impairments in the same "program"), it is important to include task force members with expertise in each of the subgroups or subprograms.

It is important that the task force not be too large to reach agreement on crucial decisions and not be too small to include the full

range of experience, expertise, and questions with which the program evaluation must deal. In the case where a very large number of programs is to be evaluated (e.g., the program for handicapped students exists on each of 106 community college campuses in California), it may be prudent to draw representatives from each of the general geographic areas covered by the programs rather than from each of the individual programs.

It is also useful to consider members' organizational position and level of professional experience. For, in addition to overseeing the evaluation, they are also in large part responsible for communicating (both formally and informally) the worth of the evaluation process and results to others in the field. If they are not supportive participants or effective communicators, success will be much more difficult to attain. The evaluation is much more likely to succeed if task force members are in a position of power within their organization and have sufficient experience and skills to put the information to use. Hence, the program representatives on the task force probably should have some administrative responsibilities as well as some responsibilities for planning or providing services or instruction. Task force members drawn from the policymaking level should probably have fiscal as well as program interests and expertise.

In addition to the above, there is a critical constellation of characteristics that all task force members should have: a positive view of evaluation in general, a flexible view of the program and possible changes which may occur as a result of the evaluation, a perception that the evaluation benefits outweigh the risks, and a commitment to use the results

to effect recommended changes (Alkin, 1985). People without evaluation experience may be used to operating from intuition rather than from data, so the duties of the task force may be new and somewhat uncomfortable to them at first. Prior experience in evaluation or research is useful in providing appreciation for the necessary technical aspects of the measures and for the feasibility of the data collection plan. In addition, recommendations on these matters from external evaluators are more likely to be appreciated and incorporated without undue delay or resistance. On the other hand, previous experience with poorly conceived or conducted evaluation efforts or those in which results were not used may have led members to distrust evaluators and to expect that the results of the evaluation will never really be utilized, especially for local program improvement. If several members fall in this category, a good deal of time may be needed to change their attitude.

Who should direct the task force? Considering the dual focus of the evaluation, the director must have several characteristics---social, political, managerial, professional and technical--- that will help ensure the success of the endeavor. It is important that all task force members, especially representatives of program personnel (who commonly feel that evaluation is a top-down imposition), view the director as "one of us." A site director of one of the programs, a peer of the other members, may fill this need well if he or she also possesses other necessary traits. Since the members will undoubtedly have divergent opinions on most issues yet be required to reach consensus, it is imperative that the director have excellent interpersonal and communication skills and have power and

political acumen. The director's effectiveness in helping the group reach consensus may reflect his or her ability to use the group dynamics to advantage, for example utilizing peer pressure on members who may create problems or stall group progress. The sheer size of the task also requires that the director have strong organization and management skills.

Due to the nature of the task, the task force director must not only have a strong background in evaluation but also be personally committed to seeing the results put to use. Research and experience suggest that evaluations are more effective when users are included in all aspects of the evaluation (Burry, Alkin, & Ruskus, 1985). This may necessitate that the director tactfully educate some task force members and local program personnel about evaluation or measurement concerns. In order to guide the task force well, the director also should have a good sense about how much information is both necessary and feasible to collect.

During at least the initial evaluation, if not the entire life of the task force, the expertise, guidance and encouragement of an external evaluation expert is likely to be needed. The evaluator must have credibility in the eyes of both the program staff and policymakers. Since the evaluator's role is to provide technical expertise and assist the task force, the evaluator must share many of the director's qualities. As with the director, it is important that the evaluator be personally commited to seeing the results put to use. The evaluation is more likely to be successful if the evaluator enjoys a good rapport with the task force members and adopts a collaborative role, in which he or she views the users as colleagues who can help guide the evaluation and who have legitimate

questions that are entitled to the evaluator's attention. The evaluator must be sensitive to the program's political dynamics, background, power or prestige. Group facilitation skills will undoubtedly be useful in resolving conflicts and negotiating issues. The evaluator must have ample experience and understanding of the technical requirements of the sorts of measures to be used as well as a realistic expectation of feasible measurement procedures in the given setting. As the task force and its mission attain comfortable age and experience, the need for an external evaluator may recede.

Step 3. Reach Consensus on Information Needs

The first duty of the task force members is to define the purpose of the evaluation and set the ground rules. Beyond agreeing that the evaluation should provide information to guide both policymaking and program improvement, the task force must deal with what balance between these two needs is to be their goal. Are both to be emphasized equally, or should one be more heavily weighted? How is the information to be used? Will it be tied to an audit or accreditation? Is it necessary or desirable for site-specific information to be kept anonymous at the policymaking level?

The second duty of the task force is to reach consensus on what specific kinds of information need to be collected. All task force members should participate in this process, drawing on input from others at the policymaking office and the individual sites. Users' information needs may be stated in terms of questions to be answered or issues to be addressed. Clarifying why users need the given information will help the group set priorities and be certain that their intentions will really meet the need precisely---not just meet it "almost."

Consensus on the general kind of information that needs to be obtained across all sites will provide a uniform information base for the two evaluation purposes. This uniform information base will be referred to here as the "core" information, indicating that it contains the basic information that is required from each program site. In addition, the task force may decide to allow individual sites to supplement the core information with additional site-specific information to be used in program improvement. If this is the case, the task force needs to reach agreement on the type and amount of information each site can collect to ensure that it will not jeopardize the success of evaluating the core issues.

Consensus on information needs is necessary because it guides all future activities of the task force and data collection efforts, particularly development of the specific measurement instruments used at every program site; hence, real, not just rhetorical, agreement by all task force members is imperative for the evaluation to be successful.

One strategy for reaching consensus is for the task force director to present a fairly comprehensive list of tentative questions or issues to which task force members may add others. Members then rank order the issues by importance. Alternatively, each issue could be rated according to its degree of importance. Results can be tallied and further discussed by the group until wholehearted agreement is reached. A good rule of thumb in reviewing all proposed issues or questions, either for the core or the optional portion of the evaluation, is: How, specifically, will this information be used to improve the program or make policy? Will this piece of information really help us make our decisions? The tendency is to say,

"That sounds important; of course we need to know that," without seriously questioning how the information will be summarized and translated into proposals for action.

Continued agreement across time may require periodic discussion and particularly clarification, if new members are added to the task force or new concerns are encountered. Significant changes in the purpose, balance, or issues addressed that occurs at any later point in the process will undoubtedly result in wasted time and effort. Hence, it is important to allow sufficient time during this step for the task force to reach a consensus with which everyone is satisfied before proceeding.

Typically, the core information needs include:

a. Who is involved?---a description of program participants

b. What does the program consist of?---a description of program services and/or instruction (which may compare what the program actually is with what it is intended to be)

c. What are the program's effects?---outcome and experiential data for all participants---students, staff and faculty---including performance and attitude data on both intended and unintended outcomes

d. What are the program's strengths and weaknesses?---conclusions and recommendations based on the other information.

Step 4. Specify Information Base and Develop Measures

After the information needs have been defined by the task force, the next step is to write precise specifications for the core information base and the site-specific issues, creating a sort of map that lays out exactly what information is to be collected and how. Specifications cover the content, type of measure, item format, and respondent or source of

information. See Figure 1 for an example.

Figure 1.
Sample Specification

CONTENT	TYPE	FORMAT	RESPONDENT/SOURCE
Overall perception of program's strength:	questionnaire	open-ended	25% random sample of all students
a. ways program has benefitted you			
b. describe one part that works well			

These specifications will guide development of the actual measures to be selected or created. Both activities, specifying information and creating or selecting measures, have been included in Step 4 because it is often easier to think of them simultaneously. Many people prefer to think in terms of sample items when considering each of the components to be specified rather than trying to consider them in the abstract. In addition, the process of creating measures often suggests new options for the components that were not previously considered, thus causing modification of the original specifications.

When specifying the information base, it is important to note the variety of measures that may be used, such as questionnaires (open-ended as well as multiple-choice), interviews, observations, inspection of written records, and performance tests. There is also variety in the sources of data, all those who participate in the program, the most important of whom

are students, faculty, staff, and administrators. Occasionally there are others affected by the program, such as teachers who work with students after they have passed through the program, from whom the task force may want to collect information. It is usually informative to use triangulation of data---asking for the same information from several different sources. Some types of measures, such as follow-up interviews with students, may best be used by trained data collectors during a visit to the program site. Others may be used by regular program participants during the course of the program activities. The task force will probably want to make full use of all these options.

Although many evaluations in the past have tended to focus on quantitative data and have eschewed using much qualitative data due to logistical problems, we would emphasize the value of both quantitative and qualitative data. However, it is easier to collect and use qualitative data if it can be partly quantified on objective, standard forms. For example, demographics on the participants, a description of program services, and a comparison between the proposed program activities and actual services provided may be gleaned from records, interviews and observations. The information may then be used to fill out standardized, highly structured forms common to all sites. An example of one such form is appended to this report.

Qualitative data on how individuals experience the program may be collected via an open-ended questionnaire intended to allow issues to emerge from comments made. (A sample questionnaire is appended to this report.) This data may then be categorized and tallied on standard forms,

thereby quantifying the major concerns expressed. If the source of the information is known, interviewers may follow-up on some of the issues with a sample of the people who made relevant comments.

Student outcome data may often be easily quantified, either through the use of existing tests (e.g. curriculum-based tests) or measures developed by the evaluator and task force specifically for this evaluation.

In many cases the desire to gather more information than is necessary or feasible given constraints of personnel, money and time will remain a problem at this stage too. The task force, therefore, may again have to prioritize the list of proposed test items, interview questions, and so forth, to obtain a reasonable amount to measure the agreed upon information base.

To save time and utilize available expertise, the evaluator can take primary responsibility for developing a tentative plan for the specifications, based on the task force's consensus on needs and priorities set in Step 3, with specific input from both policymakers and program specialists on the task force. The initial plan can be presented to the task force as a whole for review. Then the task force can select and refine those ideas which are most useful and generate additional ideas for measures to answer any remaining questions. This model, in which much of the generation of specifications (and measures) is done by the evaluator rather than by the committee as a whole simplifies the job of the task force and allows the evaluator to build in from the beginning certain necessary characteristics of a good evaluation (e.g. sound technical aspects of the measures and a realistic amount of information to be

gathered). At the same time, allowing task force members to refine and add measures gives users a sense of ownership of the measures rather than a sense of imposition by the evaluator.

The final task in Step 4 is to select or create appropriate measures (test items, interview questions, observation protocols, questionnaire items, standardized record forms, and so forth) following the specifications just developed. Again, the major responsibility for development of the measures probably best lies with the evaluation specialist, consulting the task force for specialized input, review and modification. The more precise the specifications drawn up earlier, the more straightforward is the development of measures.

Step 5. Develop Procedures for Data Collection & Report of Results

The next set of responsibilities for the task force is to oversee the development of a standardized system for using the measures developed in Step 4 to collect information and write reports for the individual program sites and for the policymakers. A system that is standardized across sites is most likely to provide the uniform information base desired for this type of evaluation.

The specific subtasks include:

a. specifying who will collect which data and when,

b. training data collectors,

c. orienting program participants to their role in the evaluation,

d. planning the reports.

A. Specifying who collects which data and when?

How the data will be collected is largely a function of what the data

will be. If the task force wants to compare program plans with actual implementation and also obtain an in-depth, qualitative picture of participants' views of the program, it will probably be necessary to visit each program site to actually see the program in action and talk directly to participants. "One-shot" data, such as performance tests, can also be collected during this visit if desired. On the other hand, if only one-shot,paper and pencil measures are to be used, a site visit may be unnecessary, and the data may be able to be collected by the local program director at each site. If the task force needs data from two or more points in time, such as pre- and post-instruction measures, a multi-phase data collection process will be necessary.

In most cases, a multi-phase process is most appropriate. First, a portion of the data is collected under the direction of the local program director at the individual site, including pre- and post-instruction measures of student accomplishment, faculty, staff and student impressions of the program via written questionnaires provided by the task force, and demographic characteristics of the students, which is available through school records.

Second, a short (1- to 3-day), well-structured site visit by a small team of trained people examines the actual implementation of the program compared to its planned intentions. The team also personally interviews a sample of the program participants. This approach allows them to follow-up on the information obtained during the first phase of data collection, which the team has examined prior to their visit. During the site visit, team members meet regularly to share impressions and cross-validate their observations.

The use of a site visit team has a couple of advantages. It provides the opportunity to bring together data on the same issues from several different perspectives---a means of data triangulation. In addition, the structured, collaborative approach allows qualitative data to be systematically collected and organized for presentation within a very short period.

Who comprises the team? Effective site visit teams utilized by Alkin & Stecher (1981) and Alkin & Ruskus (1985) consisted of three members: the team leader is a program analyst from the policymaking office; the second member is a program administrator who works at a different site from the one being evaluated; and the third member is either an external consultant or program staff member who has had experience or training in naturalistic observation. This seems to be an appropriate model, although the task force may wish to expand the team or make modifications in the members depending on the size and nature of the programs to be visited.

B. Training data collectors.

Regardless of who collects the data, it is crucial that they be given very specific directions. This is particularly true when a group of people are expected to work together as a team. A series of explicit handbooks drawn up by the task force for the site visit team members and the local program administrator can help assure consistency of data across sites. Such handbooks introduce team members to their purposes and process, describe their specific duties, provide tentative schedules and standard data forms, and provide sample eports to illustrate how data may be analyzed and synthesized into the two reports. Some sample handbook components are appended to this paper.

C. Orienting program participants.

In addition to training provided to the data collectors, there is also a need for orienting the program staff at each of the sites to be evaluated. They need to know quite a bit of information in order to cooperate and fulfill their roles: the purpose of the evaluation, how to administer any measures they may be responsible for, how large a sample to select and how to select it if measures are not being given to everyone, where and when to send the information collected, expect to hear the results. If there is to be a site visit, the program staff need to learn what to expect of the visit and should be given relevant details about scheduling and logistics. They may also be given an opportunity to request that the site visit team investigate certain concerns they may have.

D. Planning the reports.

The task force must also plan how the data will be organized into reports for the two major purposes of this evaluation: policymakers and individual program sites. These reports will probably be somewhat different due to the different needs of each of these users. Individual sites need specific information for program improvement, while policymakers may need only a general picture of the overall program, with less or no detail about individual sites. The consensus on users' information needs, reached in Step 3, should guide the content of the reports. The following outlines are to be suggestive rather than prescriptive. These lists of suggested content should be used by the task force to generate a sample site report, since in our experience, it is difficult to obtain a number of comparable reports when only an outline is provided to the authors. (See for an example Burry, 1985.)

The reports to each of the program sites may follow a common outline, such as the following:

1. An executive summary including the recommendations for specific program site

2. A description of the evaluation purposes and procedures

3. Qualitative report, including:

the setting, participants, qualitative methods used, positive findings, and areas for improvement

4. Report on program activities, including data on demographics of students, staff allocations, and planned versus implemented objectives (description, acceptability rating, and comments)

5. Appendices with all measures used

The report to the policymakers may include the following types of information:

1. Executive overview of contents

2. Description of evaluation purposes and procedures; demographic characteristics of participating program sites

3. Decription of what the program does: clientele, services, personnel

4. Program strengths (overall); optionally, the most outstanding component of each program evaluated.

5. Areas for program improvement: major themes of findings across sites; optionally, highlight specific program recommendations.

6. Appendices with all measures used

Who is responsible for writing each of these reports? The site reports may be written by the site visit team at the end of their visit. The team members may jointly contribute their insights to a unified description of findings and recommendations for improvement. Responsibility for preparation of the report (including answers to both core and optional questions) probably best rests with the team leader. The team may then personally share the results with local program staff and school administrators to use for program improvement. The qualitative portion of the report may, however, be most easily written in the week following the visit since time will be short during the visit itself. If no site visits are conducted, the task force will have to select members to be responsible for receiving, synthesizing and compiling each site's report. In either case, the task force may want to provide each site report author with the description of the evaluation purposes and general procedures, since these should be common to all sites.

The policymaker report is compiled by the director of the task force by aggregating the data from each of the individual site reports. Details on individual sites may be included or omitted depending on decisions made earlier by the task force. Obviously, this report is done after data from all sites has been collected.

One final note about the set of tasks in Step 5 is critical. Many an evaluator has felt that Murphy's Law must have been created to describe evaluations. The next step, piloting and refining measures and procedures, will be made much easier if ample time is allowed at this stage to anticipate problems and find solutions.

Step 6. Pilot and Refine Measures and Procedures

Prior to actual implementation of the evaluation plan it is important to try out the measures and procedures. Good intentions cannot guarantee effectiveness. It is highly recommended, therefore that procedures be piloted at a site not scheduled to be evaluated at this time so that the results can be used solely to improve the measures and procedures. The more similar the pilot test is to the actual circumstances of intended use, the more likely these efforts are to reveal problems that may arise. Depending on the scope of the evaluation and the extent to which the measures and procedures are untried, it may be prudent to do a two-phase pilot. During the first phase, the emphasis may be on the measures and data collection procedures. In the second phase, revised materials may be tried out and reports written to determine if the entire process were capable of generating user-oriented reports.

There are several problems which may arise and which the task force should be prepared for when piloting measures and procedures. Some participants in the evaluation do not do what they are supposed to, either because they fail to take the evaluation seriously or because they did not understand ahead of time how much time or effort it would take. Data collectors may misunderstand what they are required to do and may contaminate the results. Materials may arrive too late to be of use (e.g. pretests arriving midway through the instructional sequence). The measures may be targeted to the wrong people; they may include confusing and redundant items; and most commonly, the measures will be simply too long to effectively administer given the usual constraints.

In addition to piloting the evaluation procedures, it is also useful to try out the analyses and reporting strategies. These components are just as critical as the others to the overall success of the project. It is entirely possible that a measure has been created that looks good, works well with the people taking it, and yet provides data that is unclear or difficult to interpret. Such a problem may not be recognized until someone is forced to synthesize the findings in a report. In addition, this step will indicate whether adequate reports are likely to be produced with the planned materials, information, and procedures.

Step 7. Collect Information

Once the procedures and materials have been piloted and refined, the real evaluation can begin. At this point the task force's duty is to oversee the implementation of the plans and help to solve additional problems as they may arise.

Evaluations in general and site visits in particular sometimes evoke fear of "final certification" or loss of funding. Results will be more useful if the task force is aware of this possibility and can reassure sites about the real purposes of the evaluation. They can emphasize tha the visit is a chance to supplement other data with interviews and direct observation of the program and that it also provides an opportunity for the program staff to request a closer look at certain elements of the program. Finally, the visit allows for dialogue about findings rather than one-way communication.

Step 8. Prepare Reports

The final duty of the task force is to oversee preparation of site

reports and to prepare the final report to the policymakers. If the task force has done a good job of planning the contents of and procedures for writing these reports in Step 5 and has piloted and modified the report writing plans in Step 6, this final step should proceed without difficulty.

SUMMARY OF CRITICAL CONCERNS

1. The task force must agree on a uniform information base to meet the designated needs. In one evaluation new members were added to the task force midway through the process, and these new members questioned decisions that had been made without them. This caused the group to stall a number of times to educate these members or to rehash old ground. In addition, "agreement" seemed to have been reached, only to be rescinded later by some members who changed their minds or who had never wholeheartedly agreed in the first place but had not aired their concerns earlier. Utilize peer pressure by members who are "on track" to persuade others to contribute constructively.

2. The task force must involve the key users of the evaluation results, who should understand that the results are truly intended to be used, not just filed away somewhere. This requires a mind set among local program staff to use data rather than mere intuition for program improvement. Ultimately, the usefulness of data for improving local programs rests with the local users---their inclinations, attitudes and schemas. Are they truly trying to improve the program, or is personal survival their first priority?

3. The measures must be relevant to the agreed upon focus, technically adequate, able to produce meaningful and useful results, and be acceptable

to the users. This will require a delicate balance of input from users and professional evaluators. In a previous evaluation where measures were to be developed primarily by the task force members with guidance from the evaluation expert, the resulting measures were wildly disparate in size, scope and technical properties. Yet the evaluator felt constrained by needing task force members' ownership of the measures. Perhaps such ownership can be obtained through review rather than painstaking development of specifications.

4. Standardization of procedures and careful attention to logistical details will help assure comparable, meaningful information with fewer negative side effects for evaluation participants.

5. Realistic expectations regarding what can be accomplished given the available time, resources and personnel are critical to success. Task force members may be reminded of the importance of a few good measures, done well, with clear implications for action over a large, messy hodgepodge that is ultimately forgotten.

Conclusions

In this paper we have provided a model for evaluating programs for two purposes simultaneously: for program improvement and for policy making. This dual focus approach has two major advantages by providing a common information base for decisions at both ends of the continuum, it helps ensure that these decisions are in harmony with each other. In addition, this approach conserves time and effort of staff and students alike by making one data collection effort serve two purposes.

This user-oriented model reflects many years of evaluation experience and has been used successfully in several educational settings. It relies on a task force of evaluation users, who are open to the possibility of change and are committed to using the results. As a group they must reach consensus n their purposes, desired information base, and procedures, and then see that the plan is carried out.

The model evaluation plan presented here consists of eight steps, as follows:

1. Identify Evaluation Users and Their Needs
2. Form Task Force
3. Reach Consensus on Information Needs
4. Specify Information Base and Develop Measures
5. Develop Procedures for Data Collection and Report of Results
6. Pilot and Refine Measures and Procedures
7. Collect Information
8. Prepare Reports

While this plan was based on evaluation experiences in education, it

should also prove useful in other fields in which there is a need to make data-based decisions to improve local programs as well as to create policy that serves a number of such programs.

BIBLIOGRAPHY

Alkin, M.C. (1985). A guide for evaluation decision makers. Beverly Hills, CA: Sage Publications.

Alkin, M.C. & Ruskus, J. (1986) Systematizing short term qualitative evaluation. Paper to be presented at AERA, Spring, 1986.

Burry, J. (1985). Evaluation of martello community college's handicapped students programs and services: a prototype evaluation report. Los Angeles: Center for the Study of Evaluation, UCLA.

Burry, J., Alkin, M.C., & Ruskus, J. (1985). Organizing evaluations for use as a management tool. Studies in Educational Evaluation, 11, 131-157.

Morris, L.L. & Fitz-Gibbon, C.T. (1978) Evaluator's handbook, Beverly Hills, CA: Sage Publications.

-----(1978) How to deal with goals and objectives. Beverly Hills, CA: Sage Publications.

-----(1978) How to design a program evaluation. Beverly Hills, CA: Sage Publications.

-----(1978) How to measure program implementation. Beverly Hills, CA: Sage Publications.

-----(1978) How to measure achievement. Beverly Hills, CA: Sage Publications.

-----(1978) How to measure attitudes. Beverly Hills, CA: Sage Publications.

-----(1978) How to calculate statistics. Beverly Hills, CA: Sage Publications.

-----(1978) How to present an evaluation report. Beverly Hills, CA: Sage Publications.

Appendix A

Sample Standard Form for Quantitative/Qualitative Data

Part IV: Staff Allocations: Indicate only people for whom there is a discrepancy between proposed and actual activities

Name	FTE	Discussion of Discrepancy (Major variances by component #)	Rec #

Appendix B

Confidential Student Questionnaire

Directions: Please answer the following questions. To maintain confidentiality, please seal your completed questionnaire in the attached envelope. Sealed questionnaires will be read by review team and will not be shown to local students, staff, or administrators. Please be candid -- your comments can help to improve the project.

Thank you.

1. In what ways do you feel the project has been of greatest benefit to you?

2. Describe one part of the project here that you feel works particularly well.

Student Questionnaire (Cont'd)

Why do you think it's effective?

3. Describe one part of the project here that is in need of improvement.

Why do you think it needs to be improved (What's wrong with it)?

Please specify your major area of study _____

Please indicate services you have received from the project:
Financial aid ____ Counseling _____ Tutoring ____ Recruitment ____
Employment assistance ____ Childcare ____ Transportation ____

Your name _____ (Optional)

Appendix C

CHECKLIST

Team Leader

Prior to the Site Visit

- ☐ Identify college to be visited and select individuals to serve as Team Members B and C.
- ☐ Schedule visit and make travel arrangements.
- ☐ Prepare Site Visit Packet.
- ☐ Mail Site Visit Packet and Handbook with supplements and proper site visit schedule to Team Members B and C.
- ☐ Mail to Program Director the "Director's Guide" and multiple copies of Confidential Student and Faculty/Staff Questionnaires, and survey of student goals.
- ☐ Check with Director to assure that student goals survey has been conducted and confidential questionnaires distributed and returned. Remind Director to fill out Summary of Project Accomplishments form and Student Population form which are part of the "Director's Guide."
- ☐ Be certain that Director mails completed confidential questionnaires to Team Member C at least ten days prior to site visit.

During the Site Visit

Coordinate all activities including:

- ☐ Conduct brief team meeting before the visit to get oriented to procedures.
- ☐ Informal Introduction.
- ☐ Planning Meeting (Team Members A, B, and C meet with Director).
- ☐ Assign Interview Tasks (Team Members A and B).
- ☐ Campus Orientation (Team Members A, B, and C tour campus).
- ☐ Gather and Record Data (Team Members A and B jointly interview Director; Team Members A, B, and C separately interview other staff members and students, observe project activities, and review documents).

- [ ] Prepare for exit interview (Team Members A, B, and C prepare recommendations and compose brief description of findings).

- [ ] Conduct Exit Interview (Team Members A, B, and C with college president and administrators, Director and senior project staff.

After the Site Visit

- [ ] Prepare final report.

- [ ] Distribute final report to Director, college president, college administrator who supervises Director, to each of the other team members, and to Task Force Director (for synthesis in Report to Policymakers).

Appendix D

Tentative Schedule

(Two-Day Site Visit)

Time	Activities	Team Members
Evening Prior to Day 1	- Arrive at hotel	A+B+C
	- Team meeting to review site visit schedule and responsibilities	A+B+C
Day 1		
8:30-9:00 am	- INFORMAL INTRODUCTION - Meet with project staff - Explain purpose of site visit - Coffee	A+B+C
9:00-10:00 am	- PLANNING MEETING: - Meet with Program Director to review project data and to determine interview sources	A+B+C
	- ASSIGN TASKS	A+B
10:00-12:00 am	- CAMPUS ORIENTATION: Tour of campus and EOPS facilities. Visit with President or his designee to comment on purpose of of visit. Scheduled meeting and short interview with immediate supervisor of Program Director	A+B+C
12:00-1:00 pm	- Lunch	
1:00-2:30 pm	- GATHER AND RECORD DATA: Interview Director Interview other staff members and students	A+B C
2:30-5:00 pm	- GATHER AND RECORD DATA: Interview other staff members and students	A+B+C (separately)

Time	Activities	Team Members
8:00-9:30 pm	- Team meeting to review progress, discuss preliminary recommendations, and coordinate activities for Day 2.	A+B+C
Day 2		
8:30-11:00 am	- GATHER AND RECORD DATA: Interview staff members and students, observe project activities, and examine documents	A+B+C (separately)
11:00-12:00 am	- GATHER AND RECORD DATA: Interview Director	A+B
12:00-1:00 pm	- Luncheon meeting among team to monitor status of site visit and plan final activities	A+B+C
1:00-2:30 pm	- GATHER AND RECORD DATA: Final interviews, observations etc.	A+B+C (separately)
2:30-3:30 pm	- Team meeting to prepare final recommendations and summary of findings	A+B+C
3:30-4:30 pm	- EXIT INTERVIEW with college president and administrators, EOPS Director, and senior project staff	A+B+C
4:30 pm	- Depart	